

Mon Sep 22 11:34:33 2003

us-10-026-106e-7.rst

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2003, 18:23:55 / Search time 2374.81 Seconds  
(without alignments)  
16364.623 Million cell updates/sec

Title: US-10-026-106E-7

Perfect score: 1599  
Sequence: 1 aagccatgagcgggagccga.....acatccacgacatcgatg 1599

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2281392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database: EST.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estlin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hnc.\*  
9: gb\_estc1.\*  
10: gb\_estc2.\*  
11: gb\_hnc.\*  
12: gb\_estc3.\*  
13: gb\_estc4.\*  
14: gb\_estc5.\*  
15: em\_estfun.\*  
16: em\_eston.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_hnv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_mam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rdi.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrt.\*  
28: gb\_gss1.\*  
29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|-------|-------------|-----------|----|-------------|
| 1          | 592   | 37.0        | 1078      | 12 | BQ061188    |
| 2          | 490.6 | 30.7        | 508       | 10 | BE244935    |
| 3          | 388.4 | 24.3        | 430       | 10 | BE246345    |
| 4          | 383.4 | 24.0        | 836       | 13 | BUI51377    |

|   |    |       |      |      |    |           |                     |
|---|----|-------|------|------|----|-----------|---------------------|
| C | 5  | 381.8 | 23.9 | 991  | 12 | BQ056204  | BQ056204 AGENCOURT  |
| C | 6  | 361.6 | 22.6 | 476  | 28 | AO610898  | AO610898 HS 5105 A  |
| C | 7  | 359.8 | 22.5 | 964  | 3  | BO894374  | BO894374 AGENCOURT  |
| C | 8  | 348   | 21.8 | 541  | 28 | AO610868  | AO610868 HS 5105 A  |
| C | 9  | 338.2 | 21.2 | 358  | 13 | BUI30530  | BUI30530 UT-HP-BNO  |
| C | 10 | 330.8 | 20.7 | 640  | 14 | BY729118  | BY729118 BY729118   |
| C | 11 | 292   | 18.3 | 558  | 28 | AO634533  | AO634533 RPT-11-4   |
| C | 12 | 227.2 | 14.2 | 620  | 12 | BI104593  | BI104593 602891434  |
| C | 13 | 215.2 | 13.5 | 496  | 14 | CB713061  | CB713061 AMGNNUC:IN |
| C | 14 | 204   | 12.6 | 460  | 14 | AO772013  | AO772013 HS 5413 B  |
| C | 15 | 201.2 | 12.6 | 506  | 28 | CB738101  | CB738101 AMGNNUC:IM |
| C | 16 | 189.2 | 11.8 | 609  | 12 | BI657668  | BI657668 603281241  |
| C | 17 | 134.4 | 8.4  | 555  | 14 | W16444    | W16444 298812.r1    |
| C | 18 | 89    | 5.6  | 392  | 13 | BY093564  | BY093564 BY093564   |
| C | 19 | 56.2  | 3.5  | 663  | 12 | BI043372  | BI043372 BI043372   |
| C | 20 | 51.2  | 3.2  | 800  | 13 | BU237675  | BU237675 603411248  |
| C | 21 | 50.2  | 3.1  | 925  | 29 | CNS0091P  | AL053013 Drosophila |
| C | 22 | 47.4  | 3.0  | 391  | 14 | CB775730  | CB775730 AMGNNUC:IS |
| C | 23 | 45    | 2.8  | 290  | 10 | BF171796  | BF171796 PCL3410 M  |
| C | 24 | 44.4  | 2.8  | 641  | 10 | BE293269  | BE293269 601143880  |
| C | 25 | 44.4  | 2.8  | 642  | 12 | BM784359  | BM784359 K-EST0062  |
| C | 26 | 44.4  | 2.8  | 824  | 12 | BG821330  | BG821330 602724958  |
| C | 27 | 44.4  | 2.8  | 852  | 10 | BG468516  | BG468516 602510492  |
| C | 28 | 44.4  | 2.8  | 968  | 13 | BO420551  | BO420551 AGENCOURT  |
| C | 29 | 44.4  | 2.8  | 1000 | 13 | BO466710  | BO466710 AGENCOURT  |
| C | 30 | 44.4  | 2.8  | 1201 | 13 | BM360822  | BM360822 BX360822   |
| C | 31 | 44.4  | 2.8  | 1201 | 13 | BM360822  | BM360822 BX360822   |
| C | 32 | 44.2  | 2.8  | 1337 | 10 | BF087448  | BX421108 BX421108   |
| C | 33 | 44    | 2.8  | 1122 | 12 | BO059215  | BO059215 QV2-HT054  |
| C | 34 | 44    | 2.8  | 1201 | 13 | BK363899  | BK363899 AGENCOURT  |
| C | 35 | 44    | 2.8  | 1201 | 13 | BK381071  | BK381071 BX381071   |
| C | 36 | 43.8  | 2.7  | 341  | 29 | BZ367545  | BZ367545 140501.b   |
| C | 37 | 43.8  | 2.7  | 1135 | 13 | BQ278875  | BQ278875 AGENCOURT  |
| C | 38 | 43.4  | 2.7  | 280  | 12 | BI063605  | BI063605 PM4-UT010  |
| C | 39 | 43.2  | 2.7  | 709  | 28 | BI199792  | BI199792 602761012  |
| C | 40 | 43    | 2.7  | 837  | 29 | B2121065  | B2121065 CH230-317  |
| C | 41 | 43    | 2.7  | 1101 | 29 | CNS0201SY | AL196883 Drosophila |
| C | 42 | 43    | 2.7  | 1101 | 29 | CNS0201SY | AL196883 Drosophila |
| C | 43 | 42.8  | 2.7  | 878  | 13 | BU911485  | BU911485 AGENCOURT  |
| C | 44 | 42.4  | 2.7  | 514  | 13 | BY008068  | BY008068 BY008068   |
| C | 45 | 42.4  | 2.7  | 967  | 12 | BI108915  | BI108915 602896666  |

## ALIGNMENTS

|            |  |                         |              |        |                 |
|------------|--|-------------------------|--------------|--------|-----------------|
| RESULT 1   | BQ061188   | 1078 bp                 | RNA          | linear | EST 02-APR-2002 |
| LOCUS      | AGENCOURT 6863006  | NIH_MGC_99              | Homo sapiens | cdna   | IMAGE:5920208   |
| DEFINITION | 5' mRNA sequence.  |                         |              |        |                 |
| ACCESSION  | BQ061188   |                         |              |        |                 |
| VERSION    | BQ061188.1   | GI:19884982             |              |        |                 |
| KEYWORDS   | EST.   |                         |              |        |                 |
| SOURCE     | Homo sapiens (human)   |                         |              |        |                 |
| ORGANISM   | Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  |                         |              |        |                 |
| REFERENCE  | 1 (bases 1 to 1078)  |                         |              |        |                 |
| AUTHORS    | NIH-MGC  | http://mgs.nci.nih.gov/ |              |        |                 |
| TITLE      | National Institutes of Health, Mammalian Gene Collection (MGC)   |                         |              |        |                 |
| JOURNAL    | unpublished  |                         |              |        |                 |
| COMMENT    | Contact: Robert Straubeberg, Ph.D.<br>Email: cga@nci.nih.gov<br>Tissue Procurement: Lou Staudt<br>CDNA Library Preparation: Rubin Laboratory<br>DNA Sequencing by: Agencourt Bioscience Corporation<br>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:<br>http://image.llnl.gov<br>Plate: LHC2083 row: 1 column: 09<br>High quality sequence stop: 593. |                         |              |        |                 |

FEATURES

Location/Qualifiers

1. 1078

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5920208"

/tissue\_type="lymphoma, cell line"

/lab\_host="DH10b (phage-resistant)"

/clone\_lib="NH\_MGC\_99"

/note="Organ: lymph; Vector: pOT7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NH\_MGC library."

BASE COUNT 243 a 306 c 309 g 219 t 1 others

ORIGIN

Query Match 37.0%; Score 592; DB 12; Length 1078;  
Best Local Similarity 97.9%; Pred. No. 2,4e-126;  
Matches 619; Conservative 0; Mismatches 10; Indels 3; Gaps 2;

63 AGGAGGCGCCGCTGCGCCCTCCAGAAATGAGTGTCTCCAGAACTTCAGCT 122  
157 AAGAGGCGCCGCTGCGCCCTCCAGAAATGAGTGTCTCCAGAACTTCAGCT 216  
123 GATCTGACATGCTCCAGAGGCTTGGCAACCCAGAGATGACCTATTTTGGCC-A 181  
217 GATCTGACATGCTCCAGAGGCTTGGCAACCCAGAGATGACCTATTTTGGCCCTA 276  
182 TCAGAGCTCTCCAGCCGCTAGAGGAGTGGGCGAATGGAAGATGCGGGAGCAAGGA 241  
277 TCGAGAGCTCTCCAGCCGCTAGAGGAGTGGGCGAATGGAAGATGCGGGAGCAAGGA 336  
242 GCTGCTATGTTCTATGATGCTCTGAAAGAAAGAGCTGTGACAAAGTTCAAGGAGC 301  
337 GCTGCTATGTTCTATGATGCTCTGAAAGAAAGAGCTGTGACAAAGTTCAAGGAGC 396  
302 CGTGGGAGCGTTCTCCAGAGCTCAAGTCCCTGGTGGAGTCCGAAATACCTGGATTA 361  
397 CGTGGGAGCGTTCTCCAGAGCTCAAGTCCCTGGTGGAGTCCGAAATACCTGGATTA 456  
362 CTTTTTGAATGAGAGCGGCGCCCACTGCTGTGCTCAACCAAGAGAGATCCT 421  
457 CTTTTTGAATGAGAGCGGCGCCCACTGCTGTGCTCAACCAAGAGAGATCCT 516  
422 GAGTGCATATGCAAGTACCAAGTGCCTCCCTGCAATGCCCCACTGATCTGAAGTATGA 481  
517 GAGTGCATATGCAAGTACCAAGTGCCTCCCTGCAATGCCCCACTGATCTGAAGTATGA 576  
482 GGTGGCAATCTGGAAGAGGAGGCGGAGAAACAAGACCTATTTCAGTCACTCCCATGG 541  
577 GGTGGCAATCTGGAAGAGGAGGCGGAGAAACAAGACCTATTTCAGTCACTCCCATGG 636  
542 CCAGCCAGTCCAGATCACTCTCCAGCCAGCTCCAGAGCAACAACAGTCCCTCAGTCCAG 601  
637 CCAGCCAGTCCAGATCACTCTCCAGCCAGCTCCAGAGCAACAACAGTCCCTCAGTCCAG 696  
602 AACCATCTACAGTTCAGTGTCCCAAAATACAGCAAGTTCCTTAAGCCCACTGCTTCT 661  
697 AACCATCTACAGTTCAGTGTCCCAAAATACAGCAAGTTCCTTAAGCCCACTGCTTCT 756  
662 --GCTGAGAGTCCCAAGAGCAACTGGGCTTT 691  
757 TGTGGAGATCCCAAGAGCAACTGGGCTTT 788

DB

RESULT 2  
BE244935 508 bp mRNA linear EST 03-OCT-2001  
LOCUS BE244935

DEFINITION

TCBAP1D2669 Pediatric pre-B cell acute lymphoblastic leukemia

Baylor-HGSC project-TCHA Homo sapiens cDNA clone TCBAP2669, mRNA

sequence.

ACCESSION

BE244935

VERSION

BE244935.1

KEYWORDS

GI:9096765

SOURCE

EST

ORGANISM

Homo sapiens (human)

REFERENCE

1 (bases 1 to 508)

Wai, Y., Tsang, Y.T.M., Wei, G., Ku, J.M., Ali-Osman, J.R., Muzny, D., Bouck, J., Gibbs, R.A., and Margolin, J.F. Pediatric Leukemia cDNA Sequencing Project

TITLE

JOURNAL

Unpublished

CONTACT

Dr. Judith F. Margolin

at Baylor Children's Cancer Center and Human Genome Sequencing Center

1102 Bates, MC3-3320 Houston, TX 77030, USA

Tel: 832-824-4536

Fax: 832-825-4038

Email: clones@ccc.org

Citation: Carninci, P. and Hayashizaki, Y. High efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)

Seq primer: M3 primer.

FEATURES

source

1. 508

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="TCBAP2669"

/sex="male"

/tissue\_type="leukopheresis"

/cell\_type="pre-B cell"

/dev\_stage="pediatric 2 years"

/lab\_host="DH10b"

/clone\_lib="Pediatric pre-B cell acute lymphoblastic leukemia Baylor-HGSC project-TCHA"

/note="Vector: lambda PSB; Site\_1: BamHI; Site\_2: EcoRI; First strand cDNA was primed with an anchored XhoI-oligo(dT) primer (5'-GAGAGCTCGAGCGCGCCGAGAGAG(T)VN 3'; V=A,C,G; N=A,C,G,T) and then dg tailed. Second strand was primed with a BamHI-dc primer (5'-AGAGAGCTCGAGTCCGCGCGCGCAATTAATAT(C) 3'). Double-stranded cDNA was then digested with BamHI and XhoI and directionally cloned into the BamHI and XhoI lambda PSB vector. Library went through one round of normalization. Library was constructed by Wei Yu at RIKEN of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T, Itoh M, Nagaoka S, Sasaki, Okazaki Y, Muramatsu M, Schneider C, Hayashizaki Y. High efficiency selection of full-length cDNA by improved biotinylated cap trapper.. DNA Res 4: 1, 61-6, Feb 28, 1997)"

BASE COUNT 108 a 147 c 160 g 93 t

ORIGIN

Query Match 30.7%; Score 490.6; DB 10; Length 508;  
Best Local Similarity 99.0%; Pred. No. 5.4e-103;  
Matches 504; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

928 TCAGGAGCCCAAGCCCAAGCAAGATGAAAGAGCTTCAGAGAGCAAGAGG 987  
1 TCAGGAGCCCAAGCCCAAGCAAGATGAAAGAGCTTCAGAGAGCAAGAGG 60  
988 AGAGATGAGAGAGCAACAGAGATGAGCTTCAGAGCTTCATTAACCACTT 1047  
61 AGAGATGAGAGAGCAACAGAGATGAGCTTCAGAGCTTCATTAACCGTTT 119  
1048 CTTTCTGGGAGAGAGCAACAGAGCTTCAGAGAGCTGGAGGCTGGAGTCAAG 1107  
120 CTTTCTGGGAGAGAGCAACAGAGCTTCAGAGAGCTGGAGGCTGGAGTCAAG 179  
1108 GAGAGCCCAAGGCTCTCTGCTCCAGAGCAAGGCTCTCTGCTGGAGTTTCAAGCA 1167



## FEATURES High quality sequence stop: 522.

Location/Qualifiers

1. .836  
/organism="Homo sapiens"  
/mol\_type="rRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6304649"  
/issue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lymph; Vector: pOT87; Site: 1: XhoI; Site: 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC library."

BASE COUNT 185 a 223 c 240 g 186 t 2 others

Query Match 24.0%; Score 383.4; DB 13; Length 836;  
Best Local Similarity 99.7%; Pred. No. 4.3e-78;  
Matches 384; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 804 CTGACCTTTCTGACACACACCTTGTGCACTTTAGCCCAAGACCAAGTCC 863  
DB 541 CAGGACTTTCTGACACACACCTTGTGCACTTTAGCCCAAGACCAAGTCC 482  
QY 864 GTGAATGACTTGTCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 923  
DB 481 GTGAATGACTTGTCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 422  
QY 924 CGAGTCAGGGGCCCCAGCCCAACAGACAGATGAGAGAGAGAGAGAGAGAG 983  
DB 421 CGAGTCAGGGGCCCCAGCCCAACAGACAGATGAGAGAGAGAGAGAGAG 362  
QY 984 GAGGAGAGAGATGAGAGAGACACAGAGATGAGAGATGAGAGATGAGAG 1043  
DB 361 GAGGAGAGAGATGAGAGAGACACAGAGATGAGAGATGAGAGATGAGAG 302  
QY 1044 CTTCTTTCTCTGGGGGAGAGAGACAGAGGCTCAGAGGCTGTGGGGGTGAG 1103  
DB 301 CTTCTTTCTCTGGGGGAGAGAGACAGAGGCTCAGAGGCTGTGGGGGTGAG 242  
QY 1104 TCAGGAGAGGCCCCAGGCTCTCTGTGTCACAGAGGCTCTCTGTGGATTCTTCA 1163  
DB 241 TCAGGAGAGGCCCCAGGCTCTCTGTGTCACAGAGGCTCTCTGTGGATTCTTCA 182  
QY 1164 GACAGAGCTGGGCGACAGACTGTGG 1188  
DB 181 GACAGAGCTGGGCGACAGACTGTGG 157

RESULT 5 B0056204/c 991 bp mRNA linear EST 29-MAR-2002

LOCUS AGNCCOURT 6773374 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5808505

DEFINITION 5', mRNA sequence.  
ACCESSION B0056204  
VERSION B0056204.1 GI:19815544  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov

Tissue Procurement: Lou Staudt  
cDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LNC2052 row: k column: 02  
High quality sequence stop: 645.

## FEATURES Location/Qualifiers

1. .991  
/organism="Homo sapiens"  
/mol\_type="rRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5808505"  
/issue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_99"  
/note="Organ: lymph; Vector: pOT87; Site: 1: XhoI; Site: 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC library."

BASE COUNT 224 a 285 c 270 g 211 t 1 others

Query Match 23.9%; Score 381.8; DB 12; Length 991;  
Best Local Similarity 99.5%; Pred. No. 1.1e-77;  
Matches 383; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 804 CTGACCTTTCTGACACACACCTTGTGCACTTTAGCCCAAGACCAAGTCC 863  
DB 541 CAGGACTTTCTGACACACACCTTGTGCACTTTAGCCCAAGACCAAGTCC 482  
QY 864 GTGAATGACTTGTCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 923  
DB 481 GTGAATGACTTGTCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 422  
QY 924 CGAGTCAGGGGCCCCAGCCCAACAGACAGATGAGAGAGAGAGAGAGAG 983  
DB 421 CGAGTCAGGGGCCCCAGCCCAACAGACAGATGAGAGAGAGAGAGAGAG 362  
QY 984 GAGGAGAGAGATGAGAGAGACACAGAGATGAGAGATGAGAGATGAGAG 1043  
DB 361 GAGGAGAGAGATGAGAGAGACACAGAGATGAGAGATGAGAGATGAGAG 302  
QY 1044 CTTCTTTCTCTGGGGGAGAGAGACAGAGGCTCAGAGGCTGTGGGGGTGAG 1103  
DB 301 CTTCTTTCTCTGGGGGAGAGAGACAGAGGCTCAGAGGCTGTGGGGGTGAG 242  
QY 1104 TCAGGAGAGGCCCCAGGCTCTCTGTGTCACAGAGGCTCTCTGTGGATTCTTCA 1163  
DB 241 TCAGGAGAGGCCCCAGGCTCTCTGTGTCACAGAGGCTCTCTGTGGATTCTTCA 182  
QY 1164 GACAGAGCTGGGCGACAGACTGTGG 1188  
DB 181 GACAGAGCTGGGCGACAGACTGTGG 157

RESULT 6 A0610898/c 476 bp DNA linear GSS 15-JUN-1999

LOCUS HS\_5105\_A2\_G06\_SP6B\_RPCT-11 Human Male BAC Library Homo sapiens

DEFINITION genomic clone Plate=681 Col=12 Row=W; genomic survey sequence.  
ACCESSION A0610898  
VERSION A0610898.1 GI:5072174  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 476)  
 Mahatras G.G., Wallace J.C., Smith K., Swartzell S., Holzman T.,  
 Keller A., Shaker R., Furlong J., Young J., Zhao S., Adams M.D. and  
 Hood L.  
 Sequence-tagged connectors: A sequence approach to mapping and  
 scanning the human genome  
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
 99380589  
 10449764  
 CONTACT: Mahatras G.G., Wallace J.C., Hood L.  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887  
 Email: jwallace@u.washington.edu  
 Clones are derived from the human BAC library RPCT-11. For BAC  
 library availability, please contact Pieter de Jong  
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering/bac.htm)  
 or from Research Genetics (info@resgen.com). BAC end Web Server:  
 http://www.htsc.washington.edu  
 Plate: 681 row: M column: 12  
 Seq primer: SP6  
 Class: BAC ends  
 High quality sequence stop: 476.  
 Location/Qualifiers  
 1..476  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="plate:681 Col:12 Row:M"  
 /sex="male"  
 /clone\_lib="RPCT-11 Human Male BAC library"  
 /note="Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
 Male blood DNA was isolated from one randomly chosen donor  
 and partially digested with a combination of EcoRI and  
 EcoRI Methylase. Size selected DNA was cloned into the  
 pBAC3.6 vector at EcoRI sites"  
 137 c 128 g 115 t  
 96 a 137 c 128 g 115 t  
 BASE COUNT  
 ORIGIN  
 Query Match 22.6%; Score 361.6; DB 28; Length 476;  
 Best Local Similarity 89.9%; Pred. No. 3.8e-73;  
 Matches 410; Conservative 0; Mismatches 44; Indels 2; Gaps 2;

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 964)  
 Mahatras G.G., Wallace J.C., Smith K., Swartzell S., Holzman T.,  
 Keller A., Shaker R., Furlong J., Young J., Zhao S., Adams M.D. and  
 Hood L.  
 Sequence-tagged connectors: A sequence approach to mapping and  
 scanning the human genome  
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
 99380589  
 10449764  
 CONTACT: Mahatras G.G., Wallace J.C., Hood L.  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887  
 Email: jwallace@u.washington.edu  
 Clones are derived from the human BAC library RPCT-11. For BAC  
 library availability, please contact Pieter de Jong  
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering/bac.htm)  
 or from Research Genetics (info@resgen.com). BAC end Web Server:  
 http://www.htsc.washington.edu  
 Plate: 681 row: M column: 12  
 Seq primer: SP6  
 Class: BAC ends  
 High quality sequence stop: 510.  
 Location/Qualifiers  
 1..964  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="plate:681 Col:12 Row:M"  
 /sex="male"  
 /clone\_lib="RPCT-11 Human Male BAC library"  
 /note="Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
 Male blood DNA was isolated from one randomly chosen donor  
 and partially digested with a combination of EcoRI and  
 EcoRI Methylase. Size selected DNA was cloned into the  
 pBAC3.6 vector at EcoRI sites"  
 137 c 128 g 115 t  
 96 a 137 c 128 g 115 t  
 BASE COUNT  
 ORIGIN  
 Query Match 22.5%; Score 359.8; DB 13; Length 964;  
 Best Local Similarity 99.0%; Pred. No. 1.3e-72;  
 Matches 383; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Db 377 AAGAGAGAGAGATGAGAGAGACACAGAGATGGCGTCACTTCCAGCCCTCATTTGAC 318  
 Qy 1042 CACCTTCCTTCCTGGGCAAGACACAGGCTTCCAGGGCACTGGAGGCTGGTGGGCTG 1101  
 Db 317 CACCTTCCTTCCTGGGCAAGACACAGGCTTCCAGGGCACTGGAGGCTGGTGGGCTG 258  
 Qy 1102 ACTAGAGAGAGCCCAAGGCTCTCTGTGTCCAGAGGCTCTCTGTGTGGATTCTT 1161  
 Db 257 ACTAGAGAGAGCCCAAGGCTCTCTGTGTCCAGAGGCTCTCTGTGTGGATTCTT 198  
 Qy 1162 CAGACAGAGAGCTGGCCAGACACTGTG 1188  
 Db 197 CAGACAGAGAGCTGGCCAGACACTGTG 171  
 RESULT 8  
 A0610868/c 541 bp DNA linear GSS 15-JUN-1999  
 LOCUS HS\_5105\_A2\_D06\_SPEE\_RPCI-11 Human Male BAC library Homo sapiens  
 DEFINITION genomic clone Plate=681 Col=12 Row=G, genomic survey sequence.  
 ACCESSION A0610868  
 VERSION A0610868.1 GI:5072144  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 541)  
 Mahalaxas G.G., Wallace J.C., Smith K., Swartzell S., Holzman T.,  
 Keller A., Shaker R., Furlong J., Young J., Zhao S., Adams M.D. and  
 Hood L.  
 Sequence-tagged connectors: A sequence approach to mapping and  
 scanning the human genome  
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
 99380589  
 10449764  
 Contact: Mahalaxas G.G., Wallace J.C., Hood L.  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887  
 Email: jwallace@u.washington.edu  
 Clones are derived from the human BAC library RPCI-11. For BAC  
 library availability, please contact Peter de Jong  
 (pieterdejong.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering/bac.htm)  
 or from Research Genetics (info@resgen.com). BAC end Web Server:  
 http://www.htec.washington.edu  
 Plate: 681 row: G column: 12  
 Seg primer: SPE  
 Class: BAC ends  
 High quality sequence stop: 541.  
 Location/Qualifiers  
 1..541  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="Plate=681 Col=12 Row=G"  
 /sex="male"  
 /clone\_lib="RPCI-11 Human Male BAC library"  
 /note="Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
 Male blood DNA was isolated from one randomly chosen donor  
 and partially digested with a combination of EcoRI and  
 EcoRI Methylase. Size selected DNA was cloned into the  
 pBAC3.6 vector at EcoRI sites"  
 BASE COUNT 100 a 173 c 140 g 125 t 3 others  
 ORIGIN  
 Query Match 21.8% Score 348/ DB 28/ Length 541/  
 Best Local Similarity 87.3% Pred. No. 5,6e-70/  
 Matches 404; Conservative 0; Mismatches 56; Indels 3; Gaps 2;

Qy 821 CACACACCTGTGGCACTTTCAAGCCACAGACCAAGTCCGTGAATGACTTTTCT 880  
 Db 532 CACAGCACACCTGTGGCAAGTTCAGCCAGACCAAGT-CGGGAATGACTTTTCT 474  
 Qy 881 CTGTCCCCAAAAGAACTGACAGAGGGGTCAAGCCGACGCTCGAGTCAAGGCCACG 940  
 Db 473 CTGTCCCCAAA--GAACTGCAAGAGGGGTCAAGCCGACGCTCGAGTCAAGGCC 416  
 Qy 941 CACCCAAAGCAAGATGAGAAAGAAAGAACTTGCAGAGACCAAGAAAGAGATAGGA 1000  
 Db 415 CACACACAGCGAAGATGAGAAAGAGTCTTGCAGAGACCAAGAAAGATGAGGA 356  
 Qy 1001 GAACACAGAGATGGCGTCAAGCTTCCAGCCCTACATTGAACCACTTCTTCTG 1060  
 Db 355 GTACACANAAATGGCGTCAAGCTTCCAGCCCTACATTGAACCGCTTCTTCTG 296  
 Qy 1061 AAGACACAGGCTCCAGAGGCACTGGAGGCTGAGGAGTGAAGTCAAGAGGCGCAAG 1120  
 Db 295 AAGACACAGGCTCCAGAGGCACTGGAGGCTGAGGAGTGAAGTCAAGAGGCGCAAG 236  
 Qy 1121 TCTCTGTGTCACAGGCAAGGCTCTCTCTGTGTGGATTCTTCAAGACAGAAAGTGG 1180  
 Db 235 TGTCTGTGTCACAGGCAAGGCTCTCTCTGTGTGGATTCTTCAAGACAGAAAGTGG 176  
 Qy 1181 CACTGTGACTCTCTCTGGAGACAGGCTGGTCTCTGTGCTATTGTGCTGAGAAAGG 1240  
 Db 175 CGCTGTGACTCTCTCTGAGAACAGAGTGGTCTCTGTGCTATTGTGCTGAGAAAGG 116  
 Qy 1241 AAGCCCAAGGCGGGGTGGGATGGGACCAAGAAATCTTCCCA 1283  
 Db 115 AAGCCCAAGGCGGGGTGGTGTGGGACCAAGATCTTCCCA 73  
 RESULT 9  
 BU430530 358 bp mRNA linear EST 09-SEP-2002  
 LOCUS UI-HF-BN0-aer-b-06-0-UI\_R1 NIH\_MGC\_50 Homo sapiens cDNA clone  
 DEFINITION IMAGE:3066106 5', mRNA sequence.  
 ACCESSION BU430530  
 VERSION BU430530.1 GI:22769017  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 358)  
 NIH-MGC http://mgs.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Straubeberg, Ph.D.  
 Email: csapbs@emall.nih.gov  
 Eco RI site shown at the beginning of the sequence.  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
 cDNA Library Preparation: M.B. Soares Lab  
 cDNA Library Arrayed by: M.B. Soares Lab  
 DNA Sequencing by: M.B. Soares Lab  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. consortium/LML at:  
 www.bio.lml.gov/bdrp/image/image.html  
 Seq primer: M13 forward.  
 Location/Qualifiers  
 1..358  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3066106"  
 /tissue="lymph"  
 /cell\_type="germinal center B cells"  
 /cell\_line="MGC85"  
 /lab\_host="DHIOB (LTI)"  
 /clone\_lib="NIH\_MGC\_50"  
 /note="Vector: pYT3-Pac; Site\_1: NotI; Site\_2: Eco RI;



|            |  |            |   |        |
|------------|--|------------|---|--------|
| D6         |  | 267        | --ACGATATATCAAAAACCGGTTCGGCCACAAGTGAAGAATCTGCAGATTATCAAGGCTCTTG | 324    |
| OY         |  | 247        | TATGTTCTAATGATGTGCTCTGAAGAAGAACGCTGTAAACAAGTTCAAAGGACGGGTGC     | 306    |
| D6         |  | 325        | TGTGTCCCCTGATATGTGCTCTGAAGAAGAACCTGTATCTCCAAGTTCAAAGACAGATAC    | 384    |
| OY         |  | 307        | GGAAGGTTTTCTCCAGACTCMAAGTCCCCCTGGGTGAAGTCCGAATACTGGATTAACCTTT   | 366    |
| D6         |  | 385        | AAGCAGAGCTTCGGACAACGGACAGGTCCCAACGGGTGAAGTCCCGGTACTGTGAATACCTTT | 444    |
| OY         |  | 367        | TTGAAGTGAAGCGCGGCCCACTGTCTCTGTGCTCAACCGACGAGAGAGATCCTGAGTG      | 426    |
| D6         |  | 445        | TTGAGGTGAAGTATAGCCCGCCCACTCGGTGTCAACCCAGATGAGAAAGATCTCTAAGGG    | 504    |
| OY         |  | 427        | CCAAATGCAGATACACAGCTGCCCCCTGCATATGCCCACTGATATTGAAGTAGAGTGG      | 486    |
| D6         |  | 505        | TCAAGCTTACTCATACAGCTGCACCTTGACATCGGTGCTGAAGCTGAAGTACCAAGTGG     | 564    |
| OY         |  | 487        | CATTCTGGAAGAGGGGGCCGGAACAAGACCCTATTTCAGTCACTCCCATGGCCAGC        | 546    |
| D6         |  | 565        | TAATCTGGAAGAGGGGTCTGGAGAACAGAACCCGTGTTCTGACACTCCATATGGCCAGC     | 624    |
| OY         |  | 547        | CAGTCCAGATCACTCT  | 562    |
| D6         |  | 625        | CAGTCCAGATTCCTCT  | 640    |
| RESULT 11  |  |            |   |        |
| A0634533   |  |            |   |        |
| LOCUS      |  |            |   |        |
| DEFINITION | RPCI-11-478C4_TV RPCI-11 Homo sapiens genomic clone RPCI-11-478C4,   | 558 bp     | DNA   | linear |
| ACCESSION  | A0634533   |            |   |        |
| VERSION    | A0634533.1   | GI:5097168 |   |        |
| KEYWORDS   | GSS.   |            |   |        |
| SOURCE     | Homo sapiens (human)   |            |   |        |
| ORGANISM   | Homo sapiens   |            |   |        |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  |            |   |        |
| AUTHORS    | Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.   |            |   |        |
|            | 1. (bases 1 to 558)  |            |   |        |
|            | Zhao,S., Adams,W.D., Mierman,W., Malek,J., de Jong,P. and Venter   |            |   |        |
|            | ,J.C.  |            |   |        |
| TITLE      | Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready   |            |   |        |
| JOURNAL    | Map Building   |            |   |        |
| COMMENT    | Unpublished<br>Contact: Shaying Zhao, William Mierman, Mark Adams<br>Department of Eukaryotic Genomics<br>The Institute for Genomic Research<br>9712 Medical Center Dr., Rockville, MD 20850<br>Tel.: 301 838 0200<br>Fax: 301 838 0208<br>Email: hbeet@igf.org<br>Clones are derived from the human BAC library RPCI-11. For BAC<br>library availability, please contact Piter de Jong<br>(piteredejong.med.buffalo.edu). Clones may be purchased from<br>BACPAC Resources ( <a href="http://bacpac.med.buffalo.edu/ordering/">http://bacpac.med.buffalo.edu/ordering/</a> ) or from<br>Research Genet cs ( <a href="http://www.rtgf.org/cdb/hungen/bac_end_search/bac_end_search.html">http://www.rtgf.org/cdb/hungen/bac_end_search/bac_end_search.html</a> ).<br>Seq primer: T)<br>Class: BAC ends.<br>Location/Qualifiers<br>1..558<br>/organism="Homo sapiens"<br>/mol_type="genomic DNA"<br>/db_xref="GDB:7683219"<br>/db_xref="taxon:9606"<br>/clone="RPCI-11-478C4"<br>/sex="Male"<br>/cell_type="lymphocytes"<br>/clone_1fb="RPCI-11"<br>/notes="Vector: pBAC3.6, site_1: EcoRI, site_2: EcoRI,<br>RPCI1 Human Male BAC library" |            |   |        |

| BASE COUNT            | 113 a  | 156 c  | 180 g        | 109 t                       |  |  |
|-----------------------|--|--|--------------|-----------------------------|--|--|
| ORIGIN                |  |  |              |                             |  |  |
| Query Match           | 18.3%  | Score 292:   | DB 28:       | Length 558:                 |  |  |
| Best Local Similarity | 100.0%:  | Pred. No. 5,2e-57:   |              |                             |  |  |
| Matches 292:          | Conservative 0:  | Mismatches 0:  | Indels 0:    | Gaps 0:                     |  |  |
| QY                    | 1308   | GGTTCTGTGAAAGAGCTCCAGAAAGATACCTCTCTCTCTGGGACACTTGGGACCTTA    | 1367         |                             |  |  |
| Db                    | 1  | GGTTCTGTGAAAGAGCTCCAGAAAGATACCTCTCTCTCTGGGACACTTGGGACCTTA    | 60           |                             |  |  |
| QY                    | 1368   | CCACCGGAGCGCAATCTGGTCCCTCTGGGGGACCCCGAGTTCTCTTCAGACATGACCTTC | 1427         |                             |  |  |
| Db                    | 61   | CCACCGGAGCGCAATCTGGTCCCTCTGGGGGACCCCGAGTTCTCTTCAGACATGACCTTC | 120          |                             |  |  |
| QY                    | 1428   | TGCTGGGAAAGCAGCCCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  | 1487         |                             |  |  |
| Db                    | 121  | TGCTGGGAAAGCAGCCCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  | 180          |                             |  |  |
| QY                    | 1488   | GATCGGGGACAGCTGGGGGGCTGAGAGAGACCCAGAGAGACCGAGAGACCGAGATTTG   | 1544         |                             |  |  |
| Db                    | 181  | GATCGGGGACAGCTGGGGGGCTGAGAGAGACCCAGAGAGACCGAGAGACCGAGATTTG   | 240          |                             |  |  |
| QY                    | 1548   | GGGAGATTACATGGCCAGGTGAGCTGTCTCCCGGACATCCACCGAATCTGATG        | 1559         |                             |  |  |
| Db                    | 241  | GGGAGATTACATGGCCAGGTGAGCTGTCTCCCGGACATCCACCGAATCTGATG        | 292          |                             |  |  |
| RESULT 12             |  |  |              |                             |  |  |
| BI104593              | 620 bp   | mRNA   | linear       | EST 26-JUN-2001             |  |  |
| LOCUS                 | 602891434F1  | NCI_GCAP_Lu29  | Mus musculus | cDNA clone IMAGE:5036481 5' |  |  |
| DEFINITION            | mRNA sequence.   |  |              |                             |  |  |
| ACCESSION             | BI104593   |  |              |                             |  |  |
| VERSION               | BI104593.1   | GI:14555486  |              |                             |  |  |
| KEYWORDS              | EST.   |  |              |                             |  |  |
| SOURCE                | Mus musculus (house mouse)   |  |              |                             |  |  |
| ORGANISM              | Mus musculus   |  |              |                             |  |  |
| REFERENCE             | Eukaryote; Metazoa; Chordata; Scanata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  |  |              |                             |  |  |
| AUTHORS               | NIH-MGC <a href="http://mgs.nci.nih.gov/">http://mgs.nci.nih.gov/</a> .  |  |              |                             |  |  |
| TITLE                 | National Institutes of Health, Mammalian Gene Collection (MGC)   |  |              |                             |  |  |
| JOURNAL               | Unpublished  |  |              |                             |  |  |
| COMMENT               | Contact: Robert Strausberg, Ph.D.<br>Email: gcgabs-remail.nih.gov<br>Tissue Procurement: Gilbert Smith, Ph.D.<br>cDNA Library Preparation: Life Technologies, Inc.<br>DNA Sequencing by: Incyte Genomics, Inc.<br>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.jnl.gov">http://image.jnl.gov</a><br>Plate: LRAM1101 row: c column: 10<br>High quality sequence start: 5<br>High quality sequence stop: 556.<br>Location/Qualifiers<br>1..620<br>/organism="Mus musculus"<br>/mol_type="mRNA"<br>/strain="C57BL/6J"<br>/db_xref="taxon:10090"<br>/clone="IMAGE:5036481"<br>/tissue_type="spontaneous tumor, metastatic to mammary."<br>stem cell origin."<br>/lab host="DH10B"<br>/clone_11b="NCI_GCAP_Lu29"<br>/notes="Organ: lung; Vector: pCMV-Sport6; Site: 1: SalI; Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH" |  |              |                             |  |  |
| FEATURES              |  |  |              |                             |  |  |
| source                |  |  |              |                             |  |  |
| BASE COUNT            | 140 a  | 186 c  | 168 g        | 126 t                       |  |  |
| ORIGIN                |  |  |              |                             |  |  |





```
BASE COUNT      100 a      126 c      107 g      88 t
ORIGIN
```

| Query Match | Similarity | 12.8%   | Score 204         | DB 14      | Length 460 |
|-------------|------------|---|-------------------|------------|------------|
| Best Local  | Similarity | 74.0%   | Pred. No. 1.1e-36 |            |            |
| Matches     | 266        | Conservative  | 0                 | Mismatches | 99         |
|             |            |   |                   | Indels     | 5          |
|             |            |   |                   | Gaps       | 3          |
| QY          | 433        | CCAGTACACAGCTGCCCCCTTCATGCCCC---CACTGATCTGAATGATGAGGTGACAT    | 489               |            |            |
| Db          | 39         | CTACCTACACAGCTGCACCCCTGCATGCTCGGTCACTGGAACTGAAGTACCAAGGTGAAGT | 98                |            |            |
| QY          | 490        | TTCTGGAAGGAGGGGGGGGGGAAACAAGACCTATTTCAGTCACTCCCAATGGCCAGCCAG  | 549               |            |            |
| Db          | 99         | TTCTGGAAGGAGGACCTGAAGAACAAACCCCTGTTTCAGACACTCATATGGCCAGCCAG   | 158               |            |            |
| QY          | 550        | TTCAGATCACTTCCAGCCAGCTGCCAGCCGAAACACCACTGCTCAGTSCCAGAACCATCT  | 609               |            |            |
| Db          | 159        | TGAAGATTCTCTCCAGCAAGATGTGTGTGAGCGCCACTGCTTCAGCCCGCAAGACCTCT   | 218               |            |            |
| QY          | 610        | ACAAGTTCAGTGTCCCGAAA-TACAGCAAGTCTCTTAAGCCCACTGCTTCTTCTGTGAG   | 668               |            |            |
| Db          | 219        | ATACCTTAACCCACACAAATTAACAGCTGTCTTCAGAACCCAGCTGATCTTCTTANAG    | 278               |            |            |
| QY          | 669        | GTCCACAGAACGAACTGGGCTTTCCTGGTGTGCATGCTTGTCAATACGTGCTTAAATA    | 728               |            |            |
| Db          | 279        | GCCCCAAGGCGCAACAGGCTGTCTTGGCGGTGCTTACTTTCCTTCACTGATGAGCA      | 338               |            |            |
| QY          | 729        | ATTCCGCGAGGGGGGTGTGATCTGGAAGACCTCATGGGGAACCCCTGTGTTTCAGCGGCA  | 788               |            |            |
| Db          | 339        | ACGGCTCAACCGGGTACGGCATGGAAGAAATGATGAGGA-CCCTGGTTTCAAGTGGGAG   | 397               |            |            |
| QY          | 789        | AAGATGCCAGGGGCGCTGGAATTTTTCGAGACACACACACC                     | 828               |            |            |
| Db          | 398        | AAGATGCCCAACGATGGAATTTTTCGAAACAGACACC                         | 437               |            |            |

|            |   |
|------------|---|
| RESULT     | 15  |
| AQ772013/c |   |
| LOCUS      |   |
| DEFINITION | 506 bp DNA linear GSS-29-JUL-1999                               |
| ACCESSION  | HS_5413_B2_H01_SPEE_RPCT-11 Human Male BAC library Homo sapiens |
| VERSION    | AQ772013  |
| KEYWORDS   | Genomic clone Plater=89 Col=2 Row=, genomic survey sequence.    |
| SOURCE     | AQ772013.1 GI:5651741   |
| ORGANISM   | GSS.  |
|            | Homo sapiens (human)  |
|            | Homo sapiens  |

|         |  |
|---------|--|
| TITLE   | Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome |
| JOURNAL | Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)                                  |
| MEDLINE | 99380589   |
| PUBMED  | 10449764   |
| COMMENT | Contact: Mahairas GG, Wallace JC, Hood L   |

```

Plate: 989 row: P column: 2
Seq primer: Sp6
Class: BAC ends
High quality sequence stop: 506.
      location/qualifiers
      1..506

```

| ORIGIN                | BASE COUNT      | 111 a   | 153 c     | 97 g        | 144 t | 1 others |
|-----------------------|-----------------|---|-----------|-------------|-------|----------|
| Query Match           | 12.6%;          | Score 201.2;  | DB 28;    | Length 506; |       |          |
| Best Local Similarity | 80.9%;          | Pred. No. 5.2e-36;  |           |             |       |          |
| Matches 258;          | Conservative 0; | Mismatches 59;  | Indels 2; | Gaps 2;     |       |          |
| Qy                    | 969             | CTTCAGAGGACGAAAGAGAGAGAGAGAGAGAGACACAGAAATGCGCTCACCTTCAG      | 10226     |             |       |          |
| Db                    | 332             | CTTTGAAATGACAAAGAGATGAGAGATAG-GAGAGACACTAGATGGTGTTAAGTACCA    | 264       |             |       |          |
| Qy                    | 1029            | CCCTACATTGAAACACACTCTTTCTCTGGGCGCAAGACACAGGCTCCAGGGCCTCGAG    | 10886     |             |       |          |
| Db                    | 263             | CCGACACTTGAACCTCTGTTTATCATGGGG-AAGAGCTCCAGGCTCCAGGGCCTAGAG    | 205       |             |       |          |
| Qy                    | 1089            | GCTGTGTGGGTGACTCAAGGAGGGCCAGGGCTCTCTGTGTGCCAAGAAAGGCTCTCT     | 11486     |             |       |          |
| Db                    | 204             | TCTGTGTGGGTGGACTCAAGGAGGCAAGTGCTCAATTGTGTCCAAAGAAAGTTTATCT    | 145       |             |       |          |
| Qy                    | 1119            | GCTTGGGATTTCTGACAGAAAGCTGGGCAAGCACTGTGACTCTCTCTGGACAGGGCT     | 12086     |             |       |          |
| Db                    | 144             | GCGTGGGATTTCAACAGAAAGTTGGCCAGCAGATGTGACTCTCTCTGGACAGGGCT      | 85        |             |       |          |
| Qy                    | 1209            | GGGTCTCTGGGCTATTTTGGCTGAGAAAGGCGCAGGCGCAAGGGCCGGGTGGGGATGGGAC | 12686     |             |       |          |
| Db                    | 84              | GGGTCTCTGGGCTATTTTGGTAAAGAAAGGGGACAGTCAAGGGCCGGGTGGGNAATGGGAC | 25        |             |       |          |
| Qy                    | 1269            | CAAGATCTCTCCACACAC  | 1287      |             |       |          |
| Db                    | 24              | CAGGATCTAACCTTCCAC  | 6         |             |       |          |

Search completed: September 17, 2003, 22:13:24  
Job time : 2379.81 secs

Email: [wallace@u.washington.edu](mailto:wallace@u.washington.edu)  
Clones are derived from the human BAC library RPci-11. For BAC library availability, please contact Pieter de Jong ([pieterdejong.med.buffalo.edu](mailto:pieterdejong.med.buffalo.edu)). Clones may be purchased from BAC/RAC Resources ([http://bacpac.med.buffalo.edu/ordering\\_bac.htm](http://bacpac.med.buffalo.edu/ordering_bac.htm)) or from Research Genetics (<http://info@resgen.com>). BAC and Web Server: <http://www.htc.washington.edu>